Article manuscript: Information Display 12/99 pp.30-34

Calibrating Colorimeters for Display Measurements

S. W. Brown, Y. Zong and Y. Ohno

How red is red? Many display applications have stringent requirements for color accuracy where precision is critical, yet measuring that accuracy is more difficult than you might think. Medical and avionics displays must have reliable color to convey essential information. Color management systems must have accurate color data in order to perform effectively. The manufacturers who develop these products and the people responsible for purchasing them typically rely on results from commercial color measuring instruments: both colorimeters and spectroradiometers.

In many cases, results from these instruments are used to accept or reject display products without much attention paid to the accuracy of their measurements. If the colorimetric measurements of these displays are inaccurate, a user may accept an inadequate product or reject a perfectly good one.

In order to make effective use of any instrument, you need to know the measurement uncertainty, and this is true for colorimetry. Only then can you tell how well your color-measurement system is performing, and whether or not it can do the job you need it to do.

There is an increased appreciation of this issue in military and commercial fields; many applications now include requirements for the uncertainties for display color measurements. For example, some military specifications require expanded uncertainties (k=2) of 0.004 in chromaticity (x,y) for test equipment, while uncertainties of 0.005 or less in x,y are recommended in international standards for color measurements of cathode ray tubes (CRT) and liquid crystal displays (LCD). An expanded uncertainty of k=2 means that your measurements will agree with the true value to within the stated value 95.4 % of the time.

Instrument manufacturers typically calibrate their colorimeters. Product data sheets routinely specify uncertainty values on the order of 0.002 in x,y, and 2% in luminance (Y). A note of caution: these values may be misleading, in particular for display measurements. They are often based on measurements of an incandescent source with a color temperature of approximately 2856º K. These sources approximate the spectral power distribution of the Commission Internationale de l'Eclairage (CIE) Standard

Illuminant A, which is a common standard source used in both photometry and colorimetry.

There are problems with using just this type of source. If you measure an artifact with the same spectral distribution as the calibration source, the uncertainty may appear to be small. As long as you use the instrument to measure sources with spectral power distributions similar to that of Illuminant A, you could expect to get accurate results.

However, we typically want to measure the color and luminance of displays having spectral distributions that are quite different from Illuminant A. Comparative testing shows that the accuracy of color measuring instruments can vary significantly, depending on the spectra of the colors measured.

In the case of colorimeters utilizing three or four broadband detectors — called "tristimulus colorimeters" — the relative spectral responsivities of the individual channels to the CIE color matching functions (x-bar, y-bar, and z-bar) may not be matched well. This results in errors when the device is used to measure different spectral distributions.

In the case of array spectroradiometers, a number of factors can contribute to the accuracy of color measurements, including stray light, wavelength scale errors, and the linearity of the detector array.

No matter what type of device is used, the error analyses are complex and the uncertainties of color measuring instruments for various displays — and for different colors — are often not well known. So even if an instrument is calibrated accurately against an incandescent source, we don't know how well the instrument can subsequently measure the chromaticity and luminance of display colors.

Empirical evidence demonstrates how significant this problem can be. We took a number of colorimeters and measured the chromaticity of various colors of a display. We found that inter-instrument variations in the measurements were as large as 0.01 in x,y and 10% in Y, depending on the color and type of display being measured. These variations are an order of magnitude larger than expected from the manufacturers' specified calibration uncertainties, and imply that the uncertainty of some of the commercial instruments are at this level for display color measurements. Uncertainties in color measurements on this order are too large for many current commercial, industrial, and military applications.

In many applications, calibration approaches relying solely on incandescent sources are not adequate to meet current needs in display measurements. The National Institute of Standards and Technology (NIST) has addressed this problem by developing a calibration facility in Gaithersburg, Maryland, for colorimeters tailored to display measurements.

During a calibration, test instruments — colorimeters or spectroradiometers — are calibrated in direct comparison with a reference instrument while measuring various colors of actual displays. Upon calibration, a colorimeter can be used to measure any color of the display with a known uncertainty in chromaticity and luminance.

NIST Calibration Facility

Both source-based and detector-based methods are commonly used as transfer standards in calibrations. Unfortunately, displays are not colorimetrically stable and reproducible enough over long time periods to be used as a source for transfer standards. As a result, we have employed a detector-based calibration strategy.

The detector-based scheme starts with a reference instrument with known characteristics. This reference instrument and the instrument to be calibrated are used to measure various colors of a particular display. The differences in results determine the errors of a test instrument. Once these errors are known, they can be corrected when the test instrument is subsequently used.

In general, the errors — in both chromaticity and luminance — vary depending on the color of the display. A lookup table for the errors with different colors would be difficult to maintain. We chose a more manageable approach, deriving a correction matrix for the test instrument based on measured errors. This matrix makes it possible to correct for any color of a particular display, transferring the calibration from the NIST reference instrument to the test instrument.

In order for this strategy to work reliably, three cornerstones must be in place. We require a stable, well-characterized reference detector to be used for the display measurements, an effective matrix correction technique, and a display that is colorimetrically stable over the course of a typical calibration measurement sequence.

The NIST Reference Instrument

The reference spectroradiometer consists of imaging optics, a double-grating scanning monochromator for wavelength selection, and a photomultiplier tube (PMT) (Fig. 1). The output of the PMT is sent to a digital voltmeter and the voltage is recorded as a function of wavelength by a computer. The input optics include a depolarizer and order-sorting filters. The depolarizer reduces the polarization sensitivity of the instrument to less than 1% while order-sorting filters — placed in front of the spectrometer entrance slits — eliminate higher order grating diffraction effects.

This design results in a reference instrument with suitable characteristics (Table 1). It can measure CRT colors with an expanded uncertainty of approximately 0.001 in chromaticity and 1% in luminance.

Parameter	Value
Wavelength uncertainty	± 0.1 nm
Slit scattering function	5 ± 0.1 nm; triangular
Stray light factor	< 2 x 10^{-6}
Polarization Sensitivity	< 1%
Random Noise	< 0.2% of the peak signal

Table 1: Reference spectroradiometer characteristics.

The Four-Color Matrix Correction Method

The second cornerstone is an effective correction matrix technique that you can use to improve the accuracy of tristimulus colorimeters for color display measurements. The first step is to measure the differences in results from the reference instrument and the instrument being tested, when both measure various colors of a display. This data can then be used to create a three-by-three correction matrix. The correction matrix can then be used to transform the test instrument's tristimulus values to more closely approximate the reference values. The corrected values from the test instrument are thus expressed as linear combinations of the uncorrected values.

The trick here is to create a correction matrix that accurately transforms the values. Several different approaches have been developed to derive the correction matrix. For example, the American Society of Testing and Materials (ASTM) recommends a method that minimizes the root-mean-square (RMS) difference between measured and reference tristimulus values for several different colors of a display. The chromaticity values are then calculated from the corrected tristimulus values.

However, this method does not always work as well as expected. The ASTM method is based on the transformation of tristimulus values. It is therefore susceptible to luminance errors arising from display instabilities, and inconsistent alignment of the instruments that affect the accuracy of the corrected chromaticity values.

In contrast, measurements of chromaticity values are normally more stable and reproducible than measurements of tristimulus values. They are relative measurements and many of the sources of luminance error tend to be reduced or canceled. Consequently, a matrix correction method — the Four-Color Method — was developed that reduces the difference between measured and reference chromaticity values rather than tristimulus values. Because the new method is based on chromaticity values, it is insensitive to luminance errors and consequently tends to work better than previous methods.

A comparison of correction by various matrix methods demonstrates the differences in residual errors after correction (Fig. 2). In a test with ten colors of a display, the Four-Color Method reduced the corrected RMS chromaticity differences more than the other techniques.

Matrix correction techniques assume that the spectral distributions of the three primary colors do not change. However, there are small variations in primary spectra between displays incorporating the same type of phosphors and, frequently, large variations between displays incorporating different types of phosphors. The Four-Color Method works well (within ± 0.001 in *x,y*) for small spectral variations in the primary colors within the same types of displays, but is not effective if there are large variations between the measured display and the display used in the derivation of the correction matrix. A calibrated instrument can therefore measure any number of different displays utilizing similar phosphors to the display used for the calibration, but a separate correction matrix should be obtained for each different type of display.

Calibration Results

At NIST, we calibrated five disparate commercial instruments — including tristimulus colorimeters and array spectroradiometers — for a CRT against the NIST reference instrument. We used a broadcast-quality color CRT as the calibration source. This display demonstrated negligible chromaticity changes over the course of a comparison measurement, which took about 30 minutes. The five test instruments and the reference instrument measured ten colors of the CRT, and the chromaticity differences between the NIST reference instrument and the test instruments were recorded (Fig. 3).

We observed significant differences in the measured chromaticity values between the different units, with some instruments agreeing with the reference instrument to within ± 0.002 in *x,y* for all colors measured, and other instruments disagreeing by greater than ± 0.005 for certain colors. This is perhaps to be expected since these instruments varied greatly in design and complexity, and as a result, also in price. In general, they were all high-quality instruments; we would expect even larger differences with lower grade instruments.

Using the primary color measurements and white, we derived the correction matrix for each instrument using the Four-Color Method. We then applied the corrections to the ten colors measured, and calculated the residual chromaticity differences between the NIST reference instrument and the calibrated test instruments (Fig. 4) The results provide a dramatic illustration of this calibration method's effectiveness.

We obtained similar results for the luminance measurements, with the calibrated instruments agreeing much better with the reference instrument results for four out of the five instruments. In the one exception, the difference in luminance measurements between the test and the reference instruments was not reduced upon application of the correction matrix. This result underscores that when experimental noise or luminance errors are significantly large, the Four-Color Method is not the best suitable for luminance correction.

We know that we can measure various colors of the CRT with the reference instrument with an expanded uncertainty of 0.0012 or less in *x,y* and that the calibrated instrument agrees with the reference instrument to within 0.001 in *x,y*. Therefore, the expanded uncertainty of chromaticity measurements of the CRT using the calibrated instrument — taking the maximum uncertainty in each case and taking the root-sum-square of the uncertainties — is 0.0015 in *x,y* or less. Taking into account the repeatability of the test instrument as well, the expanded uncertainty of chromaticity measurements of a CRT using a NIST-calibrated instrument should be on the order of be 0.002 or less.

Our results demonstrate that this calibration procedure can be used to correct measurements from different colorimeters, using a reference instrument for comparison, even when those test instruments produce large differences in uncalibrated measurements for the same display colors. This calibration can bring the expanded uncertainty for these devices to 0.002 in *x,y* or less — well within the limits required for general test equipment and those recommended for international standards. The NIST facility will offer official calibration services using this technique early next year. As a result, display manufacturers and users will be able to know if the red on the screen is truly red.

Acknowledgments

We would like to thank Joe Velasquez and John Grangaard for useful discussions. This work was supported by the U. S. Air Force under CCG contract 97-422.

Author Bios:
Steven W. Brown is a Physicist, Yuqin Zong is a Guest Researcher, and Yoshi Ohno is a Physicist in the Optical Technology Division at the National Institute of Standards and Technology, 100 Bureau Drive**,** Gaithersburg, MD 20899. All correspondence should be addressed to Yoshi Ohno at 301-975-2321, or by email at ohno@nist.gov.

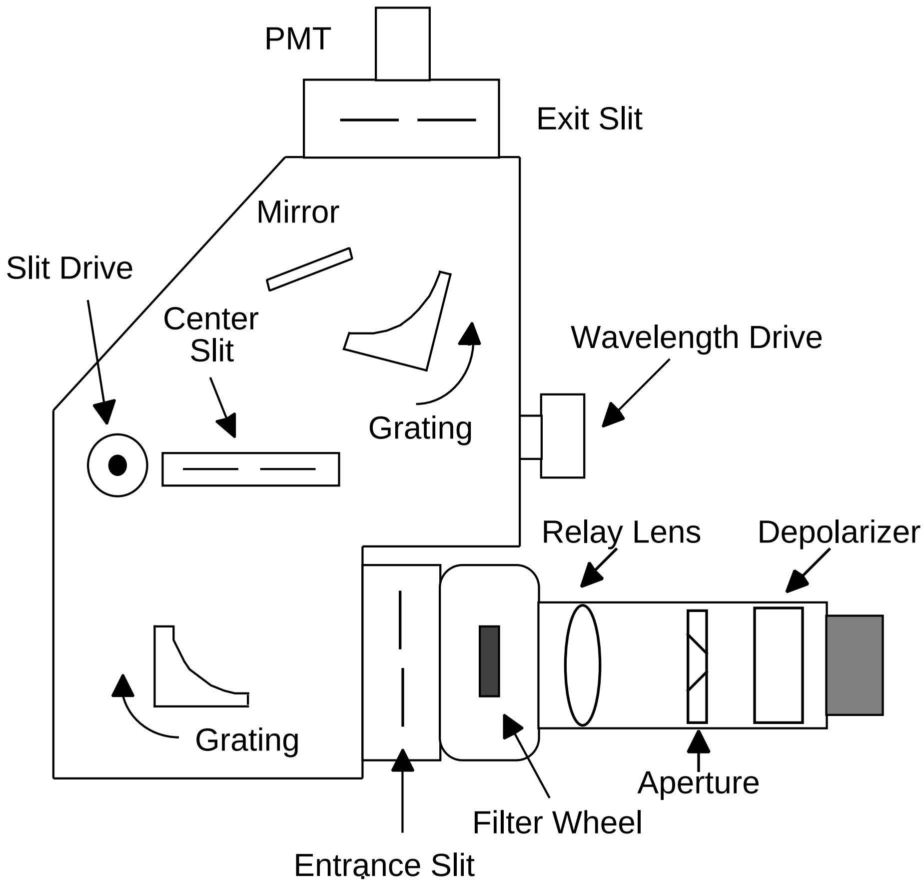

Figure 1. Schematic diagram of the NIST reference spectroradiometer.

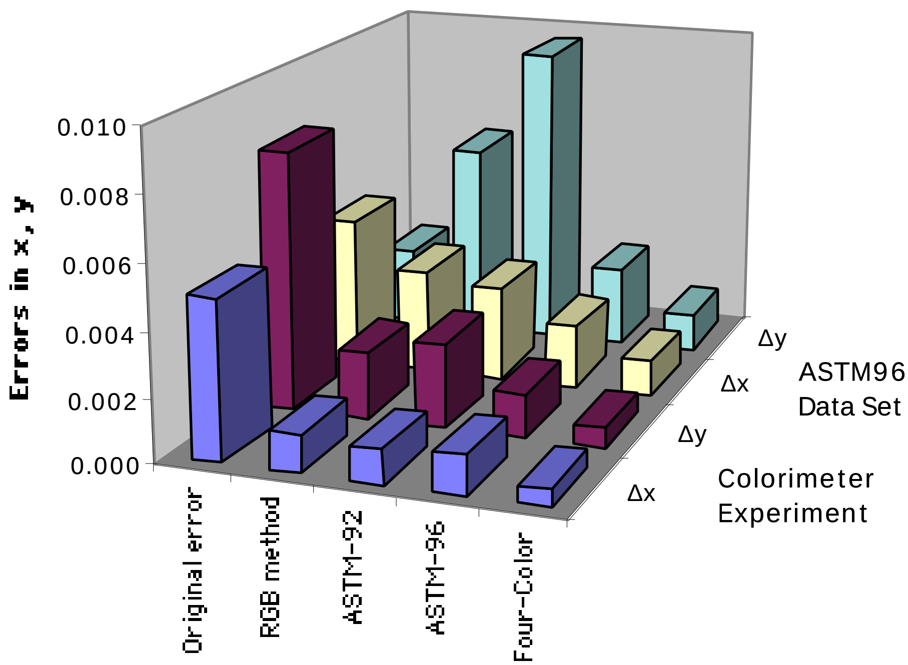

Figure 2. Root-mean-square differences in chromaticy for ten CRT colors, after correction by various matrix methods.

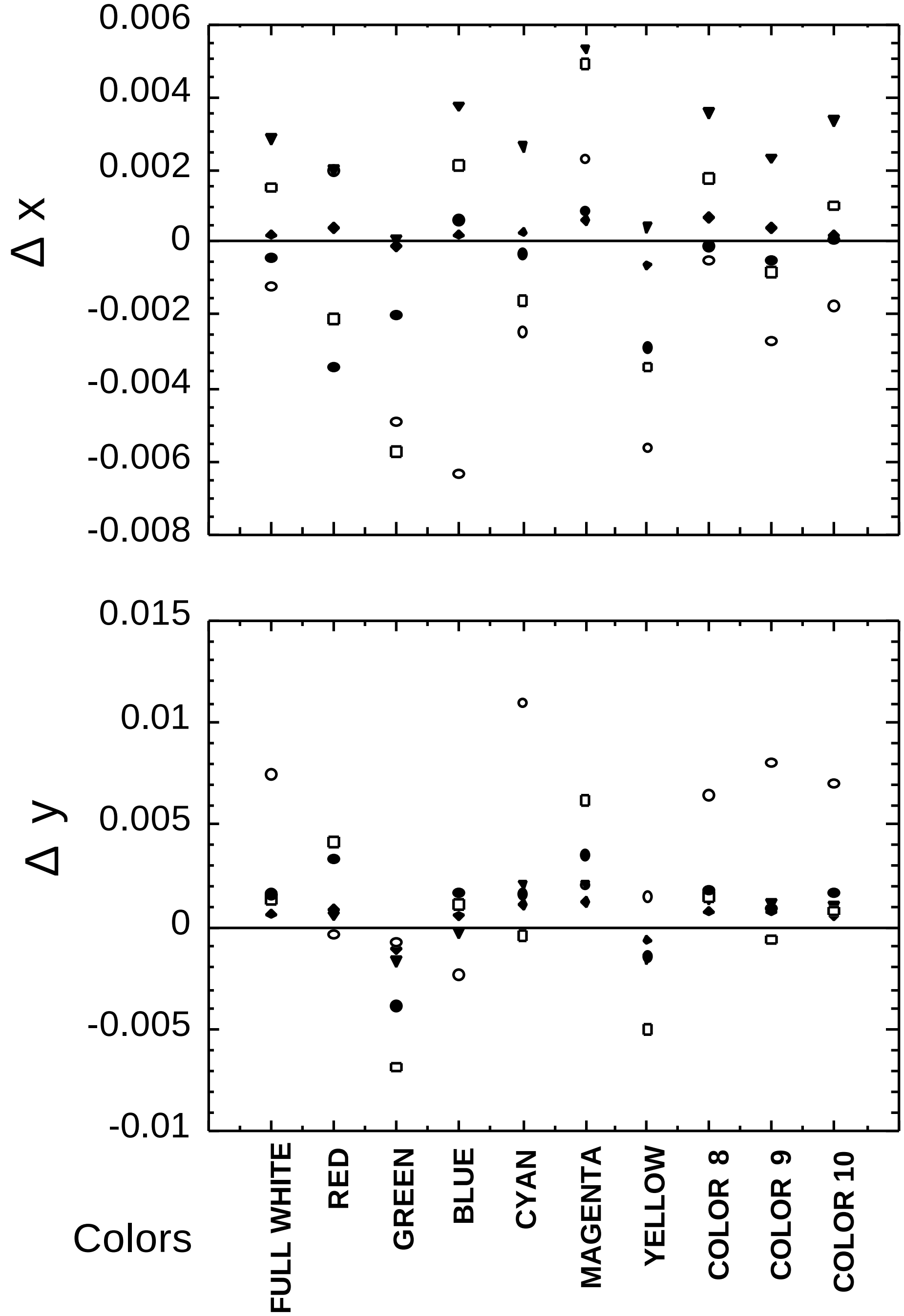

Figure 3. Uncorrected chromaticity differences between five test instruments and the NIST reference instrument for ten colors of a CRT display.

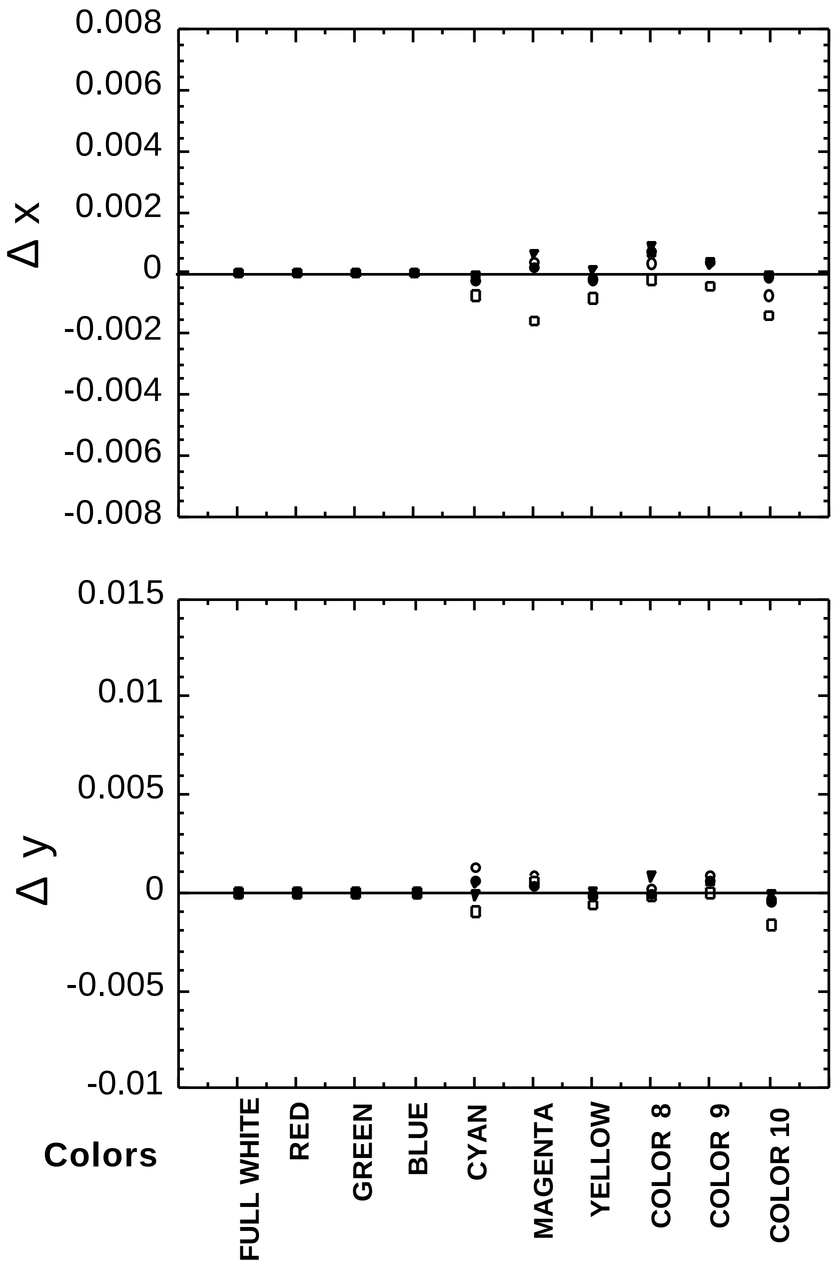

Figure 4. Residual chromaticity differences between five calibrated test instruments and the NIST reference instrument for ten colors of a CRT display.